

Lambert Sen. David

ENTERED



1600

RECEIVED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/993,192

DATE: 07/30/2003

TIME: 15:45:47

Input Set : N:\Crif3\RULE60\09993192.raw.txt

Output Set: N:\CRF4\07302003\I993192.raw

AUG 11 2003
TECH CENTER 1600/2900

of

```

1 <110> APPLICANT: Korea Research Institute of Bioscience and Biotechnology
2      Dong Kook Pharmaceutical Co.
3      Rhee, Sangki
4      Choi, Euisung
5      Kang, Hyunah
6      Sohn, Junghoon
7      Bae, Junghoon
8      Kim, Moowoong
9      Agaphonov, Michasel
10 <120> TITLE OF INVENTION: Hansenula polymorpha mutants and process for the preparation
of    recombinant proteins using the same
11
12 <130> FILE REFERENCE: 4220-116 US
13 <140> CURRENT APPLICATION NUMBER: US/09/993,192
14 <141> CURRENT FILING DATE: 2001-11-14
15 <150> PRIOR APPLICATION NUMBER: US/09/674,617A
16 <151> PRIOR FILING DATE: 2001-01-03
17 <150> PRIOR APPLICATION NUMBER: PCT/KR00/00173
18 <151> PRIOR FILING DATE: 2000-03-04
19 <160> NUMBER OF SEQ ID NOS: 9
20 <170> SOFTWARE: PatentIn version 3.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 2218
24 <212> TYPE: DNA
25 <213> ORGANISM: Hansenula polymorpha DL1
26 <300> PUBLICATION INFORMATION:
27 <308> DATABASE ACCESSION NO: U67174
28 <309> DATABASE ENTRY DATE: 1996-08-17
29 <313> RELEVANT RESIDUES: (1)..(2218)
30 <400> SEQUENCE: 1
31      ctcgagtatg attactttgt cttaattttt atttggtgta attttttact accatttttaa      60
32      tgacgtggtg tacgggtgtg aatacataaa cacctcagta gccatagcag ctaccatgga      120
33      tgatagggca cttttggaaa aggaatctga aaacaagaag aagcacggca tacggtggtc      180
34      aacagctcaa gaaaggccaa tgcgataaat ctcatTTTTa tacgaacacc caccgaagcc      240
35      cttatcggtc ttttgacgca gcagcttaat tatctgaggc tgggtatcaa tttttgcctt      300
36      ccagataaaa tatattttcc ctatttatat tccatccgtt taccctagge aatcgtccaa      360
37      aaaaacaaga attctgttat atttcaattc ccattatgaa gctctcgatg tctattctgg      420
38      ccctagtggc cactagcatt gggctggcgc aagccgctgc catcaaaaaa gataccgcgg      480
39      ggcaacaccc cctcggaatg aactccaact ttggcgaatg tttcgtggag aaaggcaagt      540
40      ctttgctcga ccgtgtttct gaggttgtga gcgagtcgtc caagaatatt tccccagaat      600
41      tgagggacat ttggaatgaa atggaggcca agttcccaga taccctcaga aacatgaagc      660
42      tcaagtcaga gccggcagtc aaaattacaa agaaacctgc cgatttttgg gacttcaatg      720
43      ttctcaatga gaagttctcc aactacaagc tgagggttaa gaagaccgac cggggagcat      780
44      tgggactgga ccacacaaga cagtactctg gatacttgga tgtggaggac gaagacaagc      840

```

RAW SEQUENCE LISTING

DATE: 07/30/2003

PATENT APPLICATION: US/09/993,192

TIME: 10:45:47

Input Set : N:\Crif3\RULE60\09993192.raw.txt

Output Set: N:\CRF4\07302003\I993192.raw

| | | | | | | | |
|----|------------|------------------------------------|-------------|-------------|------------|-------------|------|
| 45 | atttcttcta | ttggatgttt | gagtccagaa | atgacccggt | caacgaccct | gtgattctgt | 900 |
| 46 | ggctcaacgg | tggtccagga | tgctcttctt | tgactggaat | gctttttgag | ctcggtctgt | 960 |
| 47 | cttctatcgg | tccagatctc | aagccaatca | acaaccata | ttcgtggaat | tccaatgcca | 1020 |
| 48 | ctgtgatttt | ccttgaccag | cctgtcaatg | ttggattctc | gtactcttcc | aagtctgttt | 1080 |
| 49 | ctaacacggt | cgcagctggt | aaagacgtct | atgctttctt | ggagttgttc | taccagcaat | 1140 |
| 50 | tcccacactt | gctgaagaac | gacttccaca | tgcgcggcga | gtcgtacggt | ggtcattaca | 1200 |
| 51 | tcccagtgtt | tgctccgag | attctcacc | atgctgacag | atotttcaac | ctcacttcgg | 1260 |
| 52 | tgttgattgg | taacggtttg | accgacccac | ttaaccagta | ccattcttac | gagagaatgg | 1320 |
| 53 | catgctctac | tgatggtggc | tatgagcaac | cctggacgag | tctgagtgcg | aaggaatgtt | 1380 |
| 54 | tggagacctt | gcctagatgt | ttgtcattga | ttgaatcatg | ctacagctcg | cagtctgtgt | 1440 |
| 55 | tctcatgtgt | cccggcctcc | atctactgca | acaacgcaca | acttggaaca | ttccaaaaga | 1500 |
| 56 | cgggcagaaa | cgtctacgac | gttagaaaaga | tgtgcgaggg | aactctgtgc | tacaaagaca | 1560 |
| 57 | tggaatacat | tgaccaatat | ttgaaccagg | actttgtcaa | ggaaaagggt | ggcgtgagg | 1620 |
| 58 | ttgacactta | cgaagctgtg | aatctcgacg | tgaacagaaa | cttctgtttt | gctggtgatt | 1680 |
| 59 | ggatgaaacc | ttaccacaag | aacggttatca | atctgctgga | gcaaggctct | cctgtcctga | 1740 |
| 60 | tttacgcagg | agacaaggat | ttcatctgca | attggctcgg | aaaccaagcc | tggtccaatg | 1800 |
| 61 | agctcccttg | gtctggacac | gatgaattcg | agtcctccga | gctgtacaac | ctcaccttga | 1860 |
| 62 | aggatggcac | taaggctcgg | gaggctcaaga | atgctggcaa | gttcaccttt | gctagaatgt | 1920 |
| 63 | ttgatggagg | acacatggtt | ccatacgacc | agcctgagag | ctctttggct | atggtcaata | 1980 |
| 64 | gatggatagc | tggtgactac | tccttgggaa | ccaagaaata | aaaagctcat | gaaaacatta | 2040 |
| 65 | tgtcattgtt | gttcaactta | tgtatgattg | tttgtaaga | tacactatct | aaattcctcg | 2100 |
| 66 | ttcaataccc | agcctttccg | ttccttgtcc | catttacata | ttgatttgag | catcgatcgc | 2160 |
| 67 | acaaccttga | ctgttttttt | gtttgagacg | tcaatcttca | gctcgtcaag | aactgcag | 2218 |
| 69 | <210> | SEQ ID NO: 2 | | | | | |
| 70 | <211> | LENGTH: 2107 | | | | | |
| 71 | <212> | TYPE: DNA | | | | | |
| 72 | <213> | ORGANISM: Hansenula polymorpha DL1 | | | | | |
| 73 | <300> | PUBLICATION INFORMATION: | | | | | |
| 74 | <308> | DATABASE ACCESSION NO: U67173 | | | | | |
| 75 | <309> | DATABASE ENTRY DATE: 1996-08-17 | | | | | |
| 76 | <313> | RELEVANT RESIDUES: (1)..(2107) | | | | | |
| 77 | <400> | SEQUENCE: 2 | | | | | |
| 78 | gagctcagct | accttttggg | agccatggca | cgggtggcctt | gattgtacgt | tgctcctgca | 60 |
| 79 | atggtacaaa | tgggattaca | acaaaggtag | cgtcaaattct | cccgaaatag | aagcaccgcg | 120 |
| 80 | gatgagaaga | gaaagtatat | ctgcggagaa | ttacatcaaa | ctgaaaaagt | caaagtcgcc | 180 |
| 81 | tgccacaggg | tcagcattca | ttaaaatccc | aatgtccccc | catcacgttc | gcaagctgag | 240 |
| 82 | tgaaaacacg | ccactgtcgc | ctatcgactt | tcttctggac | gactatatgg | cgaaccacaa | 300 |
| 83 | ctttacaagg | gcaaaaaaca | gttcgatgag | gttttgaagc | tgtcaaacta | cggatcgata | 360 |
| 84 | tcttaatcag | atcacttgcc | tattctggaa | gccgcaccag | tactagttgg | gatagcgtaa | 420 |
| 85 | atgctcaaaa | ctgatgtggt | tatgtgtgac | gcactaccta | aaatatgagg | gtggcaaaga | 480 |
| 86 | gagtaaaatt | tgggggttcat | taagtagggt | ccataattta | caaacctgtt | ctgaatagag | 540 |
| 87 | agtcaccaat | ccagttgccc | caaaacottg | tcaaaatcat | tagcatcgtg | agcatctatt | 600 |
| 88 | ataaacaag | cgaatctcaa | agatgaaact | ctctttccca | accctttact | cgttgctctt | 660 |
| 89 | tgtgcttgga | ttggtctccc | tggccgatgc | caaggtccac | aaagctccca | tcaaaaaagc | 720 |
| 90 | tcttgacacg | gctacttacc | aggatgttac | tgtcggcgat | tatgttgagt | cgttgaaagca | 780 |
| 91 | aaagtatgtt | acaacttaca | acaagtttat | cgcgcgccaa | cagaatgacc | agcaatttat | 840 |
| 92 | ctcgatcggc | aagcgttctg | acgagagtgc | aagctcgggt | cacaacactc | ctttgaccaa | 900 |
| 93 | ctatctcaat | gccaataact | tcaccgagat | tcaattgggc | actcctggcc | aatcgttcaa | 960 |
| 94 | ggttatcctc | gacaccgggt | cctctaacct | gtgggttcct | agtagcgatt | gcacgtcctt | 1020 |

RAW SEQUENCE LISTING

DATE: 07/30/2003

PATENT APPLICATION: US/09/993,192

TIME: 10:45:47

Input Set : N:\Crf3\RULE60\09993192.raw.txt

Output Set: N:\CRF4\07302003\I993192.raw

```

95      ggctgttac ttgcatacta aatacgaacca cgaatgagtcg tccacttacc agaagaacgg      1080
96      ttcttcgttt gctattcaat atggctcggg ttccctcgag ggttacgttt cccaagacac      1140
97      actgactatt ggtgaccttg tcatccctaa acaggatttc gccgaggcta ctacgagacc      1200
98      tggttgggt tttgcttttg gaaagtttga cggatttctg ggtctggctt acgacacgat      1260
99      ctcggtcaac agaatcggtc ctccaattta caatgctatc aatttgggat tgctggacac      1320
100     cccacaattc ggattttacc ttggtgacac ttogaagtcc gagcaggatg gaggagaggg      1380
101     tacctttggt ggatacgatg tgtctaagta cacaggcgat atcacctggt tgccagtcag      1440
102     aagaaaggct tattgggaag tgaagtttag tggatgccc ttggtgacga atacgtccca      1500
103     ttggagaaca ccgggagctg ccattgatac cggaacctct ttgattgctc ttccatctca      1560
104     attggtcgag attttgaact ctcaaattgg tgccgagaag tcatggtctg gacagtacca      1620
105     gatcgattgt gacaagagag actcactgcc tgacctcact ttcaacttcg atggttataa      1680
106     ctttaccatc tcgccttatg actacacttt ggaagtttct ggttcttcta tttctgcatt      1740
107     cactccaatg gacctccag ccccaattgg tccaattggc atcattggtg acgctttcct      1800
108     cagaagatac tactctgtct atgaccttgg cagagacgct gtcggatttg ctaaggctgt      1860
109     gtaatgcaag gcagttacgt attagtttca acttgctagc atgaatttct atcgggtacg      1920
110     gttgcttgac aacgattttc tttgtgtctc tattactgct ataaaacatg aaaagattgg      1980
111     attttttga cagtttctag taattgctta aatctgctct tttccttttg ttgatcccg      2040
112     gctgctctgt gtgcttctt tttttatcta tattattaga ttgtctcttc tttagcagcg      2100
113     gaagctt                                     2107
115 <210> SEQ ID NO: 3
116 <211> LENGTH: 3550
117 <212> TYPE: DNA
118 <213> ORGANISM: Hansenula polymorpha DL1
119 <400> SEQUENCE: 3
120     gaattctata ggaatgggtgc aaaacgaagg atcgtgtgga ctcatgaatg tgctatatatt      60
121     gtggctaaaa agcctgccag agctcgcagg agataagagg tctggcgctg tgctgcgcaa      120
122     aatgatcgcc ttgataaaat tcctcaatct cgatcatgaa tcagggttca agtatttttg      180
123     tttttcttt caagacgagt atgaactgga acacgctctc gagtccctaga cgaccaggac      240
124     atcaagcagg cggccgcact cgaagatctc gccgcttttg atttccaaca ctttgacgtc      300
125     gtggaggttg accagacct ttctgatgcc tatgagctcg agctcggctc cgaacagggt      360
126     tcctggttgg aacgtgaatt tcaggttgtt ggtaagagaa atcagcccca ctttgcgtgaa      420
127     ctctgttgg aaagcccgca ccagctgggt gaggccgtaa agcgggacgg gatctgcctt      480
128     gttggcgctc tggatcggaa ttggtggcac agatttttctg tattgtttgt tgatcgatga      540
129     ttgaatgtac agtatgagcg ataacgagtc gttcgacttg ttgataaatt gtagcttcca      600
130     tttgaacacc tcgccgagct tgactttagt ctggcccgaa actgaaatgc tgaggttctt      660
131     cttggagtggt ttgagagtca ccgacgagtt ggatcgtagt ttatttttct taggcaaatt      720
132     cagagcagag ttggtcgact tggattttaa tccccgaga ctaggcgacg cgattggggg      780
133     cacggtgggg gccgccataa aatccacatt tgtagtccat cgtgtcacga ttttctatt      840
134     cccgttgacc gtgcaatgga tggtgacaaa aaccggctta acccacttga tgtcgttggt      900
135     tgtagtctg tacgtgagcg acagaataga atggttctgc aacaccaatg gatactttac      960
136     caccgagtag ccggcacaca gacagtttct gatgtccata tctaccgact cgattctcac      1020
137     cgacacttct gacgcccgga cctgtccaga aagagccact tccaaagagg ccatggtttt      1080
138     ctccgtcttg atggtcttga accgcactct gaatacgggg ctgatccgca gcgacgtgtc      1140
139     ggccgttacg agttcgtaaa ggtccacatg cttcgtcagc acagtgttgc ctatagtgc      1200
140     tgtatcctgc aactctaaat tttctatctt tctcgactgt tgtcgaaacg aggttttgat      1260
141     ctgcaattga tttgaaagtt tgtcaggcac cagggtcgtt tccagtatcc atacggcctc      1320
142     agacctatg acgacttttg aatcgagcac ttttctcgct gtgttttgca agttgtatac      1380
143     caggtcgacc tcgacctgtg tcacagcga cggagtgtac accacgactt ggatgtcctc      1440
144     gtcgaaatag gcaatgtcca gcgtgtcgca ctggcgacg tgttccacgg agcccagcac      1500

```

RAW SEQUENCE LISTING

DATE: 07/30/2003

PATENT APPLICATION: US/09/993,192

TIME: 10:45:47

Input Set : N:\Crf3\RULE60\09993192.raw.txt

Output Set: N:\CRF4\07302003\I993192.raw

```

145 atcagcagaa acccggtgtt tcggaacaag cacaccaaac attggaacaa acacttatcg 1560
146 acgcgggtgt gtttaaaaaa taatattaga tcgactgctc ctactctttt aaattcttta 1620
147 tttatccag gcttagacag cacagagcta accatgagtc ggtatTTTTT actagcgtgc 1680
148 acattggctc tgcaatgtgt cgcggcctcc caggaggact acatagtcaa ggatttgcct 1740
149 ggactctcga atattcctgc cgtggtgagg ccagtgatgc atgccggaca ccttgaaata 1800
150 gatgaggaac acaacaccga gctgtttttc tggcgattcc agaatccgaa gaacaacggg 1860
151 acacaccaa cgctccaccg caatgagctg atcgtctggc tgaacgggtg ccctggctgc 1920
152 tcgtcgatgg atggagccat gatgaaaca ggccccctgc ggggtgtcaga caagttggag 1980
153 gtggagctca acccggtatc ctggacacag gtggccgata tcttgtttgt ggatcagccg 2040
154 gccggaattt tgagaaagtg gagtgcgtca aaatattcca tagtattttg gctgccagta 2100
155 gagacgaaac caagcctgcc aaagaacaat gtgtcaacat gtacgactac cgcaaacatg 2160
156 attacttccc tgcattgcga tccaattggc cagaaggcct gccacccgtg actaaatttc 2220
157 tgaatctgga tgcgtgcaa aaagccctga acctgaaatc ggcgaagaga tggcacgagt 2280
158 gtgacggaaa agtcgaattc tttttccagc cagagcactc tgtcaagtcg ttcgacctgc 2340
159 tgccaaaact actggagaaa atgaaaatcg cgtgttttgc tggcgacaag gacattatct 2400
160 gcaaccacaa atcaatagaa atggtgattg aaaaattgca aattacacca ggacaattcg 2460
161 gtttcacaaa cagcagaaag tctggctgga tctacgacgg gcaggaagtg ggtgaggttg 2520
162 agacacagtc gaacctcacg tatatcaagg tgttcaattc gtcgcataat gtgccttacg 2580
163 acctgccgga ggtgagcaga ggactattcg atataatcac aaactctatt gaaaaacggg 2640
164 cgacagacat tgtgacgcc gtttatgaca gcagaggcaa ctacaagttt gtggaggaga 2700
165 agcaggacac cgaccagaac gaggaagaag agaaggagaa acctcccaag caccatcata 2760
166 gtctcacgtt ttactggca gaggtggcga ttctggcagt gctagcatat ctgctttaca 2820
167 gcttctacaa atcgttcgcc aaatcgcgta agtctgcatt tttgtcactc tcttccaaga 2880
168 agaagaagaa gcaggtgcac tggtttgacg agagcgacat aggaatggac caggaagctg 2940
169 gcgaggcaga tcataagcct aagagcatgc tggagtgggt gttcaacaag ctgggctatg 3000
170 gaggacagta tgacacggta caggacggcc gtgacatcga gatggcgccg gtcgaagaac 3060
171 acgaagacca atttattatt caaagcgacg aggaagagtt tggccacaga tagagtacga 3120
172 ttatataata tccacgaaaa attagtttcc aattatttcc ctcttgattg taatgtctca 3180
173 aagttggatc ctctggaagc ggtccctgat taccggcggt ggaattattg gaagcgggtg 3240
174 gctactgtac aaattcacga ctctacoga ggaacagttg atcgctaaat tatccccga 3300
175 gctgagagca gactatgaac gcaacaaaga attgcggagg aaagaacagg agatgttgat 3360
176 ggagatcgtc aaacagacgg cagcgtcgaa cgatccggtg tggagacag ggccgattgt 3420
177 gtcgccgtgg gaccgggact ttacgccatc gagggaaagt ttattagtga agagagagcg 3480
178 atttgagaaa gagcaggcgg agaaaaaaca gcgcgaagag ctcgagcgtc tgaaagccga 3540
179 ggccaagctt

```

```

181 <210> SEQ ID NO: 4
182 <211> LENGTH: 17
183 <212> TYPE: DNA
184 <213> ORGANISM: Artificial Sequence
185 <220> FEATURE:
186 <223> OTHER INFORMATION: PCR primer for S. cerevisiae PRC1 gene
187 <400> SEQUENCE: 4
188 atgaaagcat tcaccag
190 <210> SEQ ID NO: 5
191 <211> LENGTH: 17
192 <212> TYPE: DNA
193 <213> ORGANISM: Artificial Sequence
194 <220> FEATURE:
195 <223> OTHER INFORMATION: PCR primer for S. cerevisiae PRC1 gene

```

17

RAW SEQUENCE LISTING

DATE: 07/30/2003

PATENT APPLICATION: US/09/993,192

TIME: 10:45:47

Input Set : N:\Crf3\RULE60\09993192.raw.txt

Output Set: N:\CRF4\07302003\I993192.raw

```

196 <400> SEQUENCE: 5
197      ttataaggag aaaccac                                     17
199 <210> SEQ ID NO: 6
200 <211> LENGTH: 17
201 <212> TYPE: DNA
202 <213> ORGANISM: Artificial Sequence
203 <220> FEATURE:
204 <223> OTHER INFORMATION: PCR Primer for S. cerevisiae PEP4
205 <400> SEQUENCE: 6
206      atgttcagct tgaaagc                                     17
208 <210> SEQ ID NO: 7
209 <211> LENGTH: 16
210 <212> TYPE: DNA
211 <213> ORGANISM: Artificial Sequence
212 <220> FEATURE:
213 <223> OTHER INFORMATION: PCR primer for S. cerevisiae
214 <400> SEQUENCE: 7
215      tcaaattcgt ttggcc                                     16
217 <210> SEQ ID NO: 8
218 <211> LENGTH: 27
219 <212> TYPE: DNA
220 <213> ORGANISM: Artificial Sequence
221 <220> FEATURE:
222 <223> OTHER INFORMATION: PCR primer for H. polymorpha KEX1
223 <400> SEQUENCE: 8
224      tggysaacg ghccwgggtg ytcbtcb                         27
226 <210> SEQ ID NO: 9
227 <211> LENGTH: 30
228 <212> TYPE: DNA
229 <213> ORGANISM: Artificial Sequence
230 <220> FEATURE:
231 <223> OTHER INFORMATION: PCR primer for H. polymorpha KEX1
232 <400> SEQUENCE: 9
233      wggratgtay tgwccrgcgt avgactcdcc                     30

```

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 07/30/2003

PATENT APPLICATION: US/09/993,192

TIME: 10:45:48

Input Set : N:\Crf3\RULE60\09993192.raw.txt

Output Set: N:\CRF4\07302003\I993192.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 10

VERIFICATION SUMMARY

DATE: 07/30/2003

PATENT APPLICATION: US/09/993,192

TIME: 10:45:48

Input Set : N:\Crf3\RULE60\09993192.raw.txt

Output Set: N:\CRF4\07302003\I993192.raw